

## BIROn - Birkbeck Institutional Research Online

Gooptu, Bibek and Miranda, E. and Nobeli, Irene and Mallya, M. and Purkiss, Andrew G. and Leigh Brown, S.C. and Summers, C. and Phillips, R.L. and Lomas, D.A. and Barrett, Tracey E. (2009) Crystallographic and cellular characterisation of two mechanisms stabilising the native fold of 1-Antitrypsin: implications for disease and drug design. *Journal of Molecular Biology* 387 (4), pp. 857-868. ISSN 0022-2836.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/1058/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>  
contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

or alternatively

# Crystallographic and Cellular Characterisation of Two Mechanisms Stabilising the Native Fold of $\alpha_1$ -Antitrypsin: Implications for Disease and Drug Design

Bibek Gooptu<sup>1\*</sup>, Elena Miranda<sup>2</sup>, Irene Nobeli<sup>1</sup>, Meera Malliya<sup>2</sup>, Andrew Purkiss<sup>1</sup>, Sarah C. Leigh Brown<sup>2</sup>, Charlotte Summers<sup>2</sup>, Russell L. Phillips<sup>2</sup>, David A. Lomas<sup>2†</sup> and Tracey E. Barrett<sup>1†</sup>

<sup>1</sup>School of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK

<sup>2</sup>Department of Medicine, University of Cambridge, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 0XY, UK

Received 16 January 2009;  
accepted 31 January 2009  
Available online  
14 February 2009

The common Z mutant (Glu342Lys) of  $\alpha_1$ -antitrypsin results in the formation of polymers that are retained within hepatocytes. This causes liver disease whilst the plasma deficiency of an important proteinase inhibitor predisposes to emphysema. The Thr114Phe and Gly117Phe mutations border a surface cavity identified as a target for rational drug design. These mutations preserve inhibitory activity but reduce the polymerisation of wild-type native  $\alpha_1$ -antitrypsin *in vitro* and increase secretion in a *Xenopus* oocyte model of disease. To understand these effects, we have crystallised both mutants and solved their structures. The 2.2 Å structure of Thr114Phe  $\alpha_1$ -antitrypsin demonstrates that the effects of the mutation are mediated entirely by well-defined partial cavity blockade and allows *in silico* screening of fragments capable of mimicking the effects of the mutation. The Gly117Phe mutation operates differently, repacking aromatic side chains in the helix F– $\beta$ -sheet A interface to induce a half-turn downward shift of the adjacent F helix. We have further characterised the effects of these two mutations in combination with the Z mutation in a eukaryotic cell model of disease. Both mutations increase the secretion of Z  $\alpha_1$ -antitrypsin in the native conformation, but the double mutants remain more polymerogenic than the wild-type (M) protein. Taken together, these data support different mechanisms by which the Thr114Phe and Gly117Phe mutations stabilise the native fold of  $\alpha_1$ -antitrypsin and increase secretion of monomeric protein in cell models of disease.

© 2009 Elsevier Ltd. Open access under [CC BY license](#).

**Keywords:** serpin;  $\alpha_1$ -antitrypsin; polymerisation; rational drug design; conformational disease

Edited by M. Guss

## Introduction

The plasma protein  $\alpha_1$ -antitrypsin is principally synthesised by hepatocytes and protects lung tissue from degradation by inhibiting neutrophil elastase.<sup>1</sup> It is the archetypal member of the serpin

(serine protease inhibitor) superfamily of proteins.<sup>2</sup> In common with other inhibitory serpins, the native fold of  $\alpha_1$ -antitrypsin is metastable and this property is utilised in the characteristic serpin mechanism of action.<sup>3</sup> The reactive loop of the molecule is oriented into solution as an ideal substrate for its target protease. Following docking, elastase cleaves the P1–P1'<sup>4</sup> bond of  $\alpha_1$ -antitrypsin, allowing the P1–P14 residues of the reactive loop to insert as an additional strand within  $\beta$ -sheet A. The enzyme remains bound to the P1 residue in a catalytic intermediate state and is translocated 70 Å from one pole of the molecule to the other.<sup>5</sup> The energy released by the formation of the new  $\beta$ -strand disables the catalytic site and stabilises the

\*Corresponding author. E-mail address: [b.gooptu@mail.cryst.bbk.ac.uk](mailto:b.gooptu@mail.cryst.bbk.ac.uk).

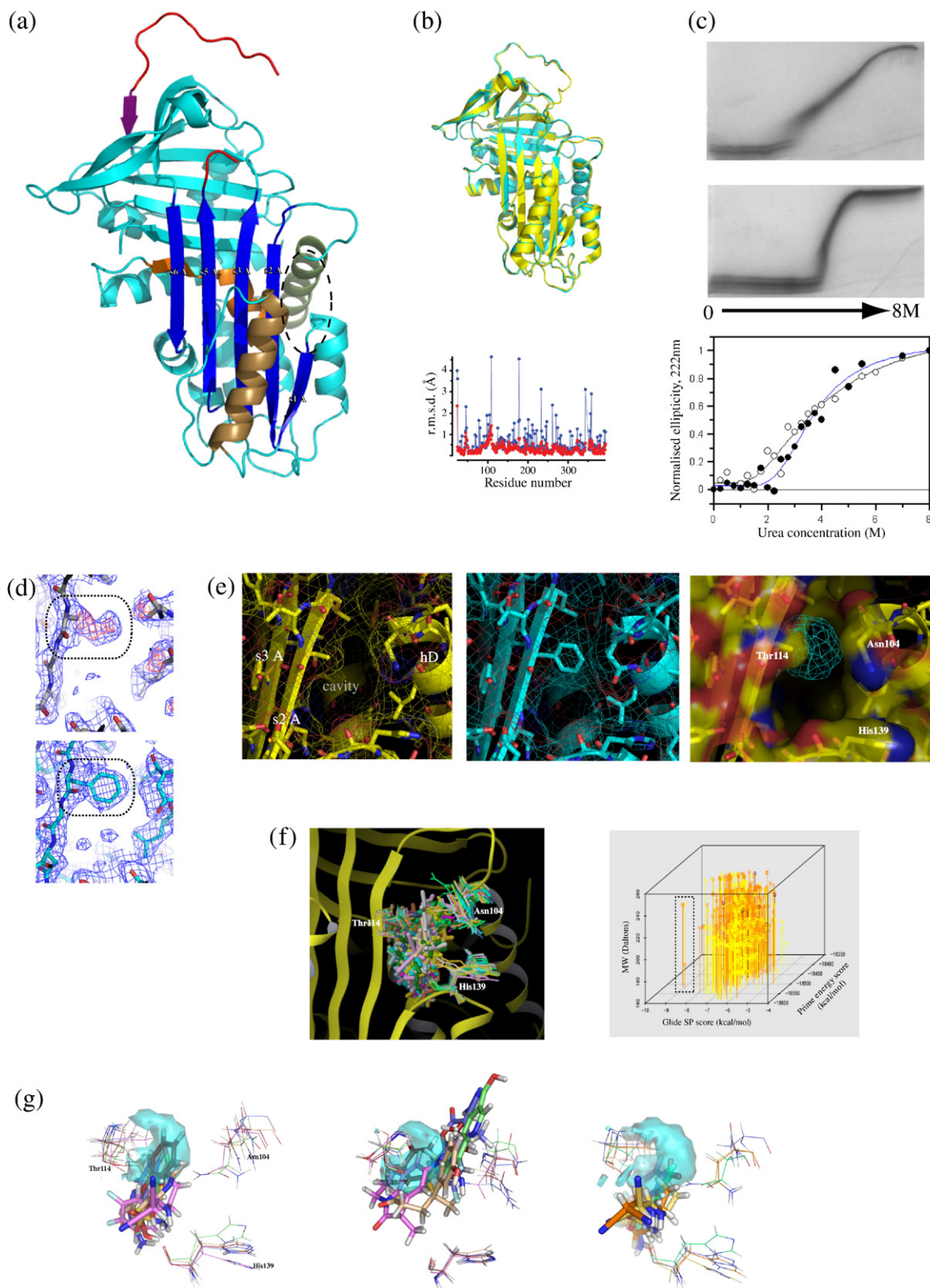
† D.A.L. and T.E.B. are joint senior authors.

Abbreviations used: serpin, serine protease inhibitor; PDB, Protein Data Bank; NCS noncrystallographic symmetry; TUG, transverse urea gradient; PEG, polyethylene glycol.

covalently bound complex, which is then rapidly cleared from the circulation.<sup>6</sup>

This dynamic mechanism renders  $\alpha_1$ -antitrypsin vulnerable to point mutations, predisposing it to alternative, pathogenic, conformational transitions—

in particular self-association into polymer chains.<sup>7</sup> These accumulate within the endoplasmic reticulum of hepatocytes rather than enter the circulation. The most clinically significant of these is the Z mutation (Glu342Lys) carried by 1 in 27 of North European



**Fig. 1** (legend on next page)

populations. Homozygotes for the Z allele have deficiency of circulating  $\alpha_1$ -antitrypsin with levels that are 10–15% of those found in homozygotes for M (wild type)  $\alpha_1$ -antitrypsin. Disease ensues from both the toxic effects of polymers aggregated within hepatocytes (neonatal hepatitis, cirrhosis and hepatocellular carcinoma) and the loss of antiprotease activity within the lung (emphysema).

Over the last 15 years, biochemical, biophysical and structural studies have defined the general scheme of  $\alpha_1$ -antitrypsin polymerisation.<sup>7–11</sup> This proceeds via an initial monomeric intermediate termed M\* in a conformation that is activated for the formation of polymers. The existence of this initiating species is indicated by fluorescence studies of  $\alpha_1$ -antitrypsin polymerisation kinetics and by the structure of the thermodynamically stabilised yet highly polymerogenic  $\delta$  conformer of the closely related protein,  $\alpha_1$ -antichymotrypsin.<sup>8–12</sup> These studies support the hypothesis that the native-to-M\* transition results from dysregulated opening of  $\beta$ -sheet A and allows the reactive site loop of a neighbouring molecule to insert aberrantly as an extra  $\beta$ -strand. Successive intermolecular loop–sheet interactions lengthen the polymer chain. Recently, an alternative model has been proposed, based upon the structure of a self-terminating dimer of the serpin antithrombin,<sup>13</sup> emphasising polymerisation

as an alternative endpoint on the folding pathway. The repeating subunit of the proposed polymer is considerably unfolded compared to the native structure with intermolecular linkage occurring via a  $\beta$ -hairpin formed from the reactive loop and a contiguous  $\beta$ -strand.

Both proposed mechanisms of polymerisation share key characteristics that can be exploited for the purposes of rational drug design. In both cases, increasing stability of the native fold will reduce the population of the unstable intermediate. Moreover, in both cases, polymerisation requires expansion of  $\beta$ -sheet A and the formation of a six-stranded sheet. For  $\alpha_1$ -antitrypsin, this involves the sealing of a solvent-accessible cavity flanking  $\beta$ -sheet A.<sup>14</sup> This, therefore, represents an attractive target for rational drug design as small molecules binding here will impede expansion of  $\beta$ -sheet A and, thus, prevent polymerisation by an allosteric mechanism. Evidence for this comes from the effects of mutating either of two residues bordering the cavity—Thr114 and Gly117—to phenylalanine. Both mutations reduce polymerisation of  $\alpha_1$ -antitrypsin by an order of magnitude *in vitro* whilst leaving inhibitory activity unaffected.<sup>14</sup> Moreover, both mutations partially rescue the defect in the total levels of secreted Z  $\alpha_1$ -antitrypsin in a eukaryotic cell model of disease. The importance of this cavity is

**Fig. 1.** Structure of Thr114Phe  $\alpha_1$ -antitrypsin and *in silico* fragment screen. (a) Crystal structure of Thr114Phe  $\alpha_1$ -antitrypsin shown in cyan with key structural features coloured. The reactive loop is depicted in red,  $\beta$ -sheet A in dark blue, the shutter region in orange, helix D in green, helix F in bronze and strand 1 of  $\beta$ -sheet C (s1C) in purple. The cavity flanking  $\beta$ -sheet A is indicated (ellipse). (b) Superposition of wild-type (PDB code: 1QLP,<sup>16</sup> yellow) and Thr114Phe (cyan)  $\alpha_1$ -antitrypsin with plot of the r.m.s.d. between the main chains (red) and side chains (blue). (c) Increased stability of the native fold in the presence of the Thr114Phe mutation. TUG-PAGE of wild-type (upper panel) and Thr114Phe (middle panel)  $\alpha_1$ -antitrypsin indicates that the mutation confers increased stability of the native fold to increasing urea concentration. CD spectroscopy allows urea stability to be compared more directly. Ellipticity values measured at 222 nm with varying urea concentrations are shown for wild-type and Thr114Phe  $\alpha_1$ -antitrypsin ( $n=3$ ). Values are normalised to allow direct comparison of the loss of initial signal. Curves were fitted in Graft 3.0 (Erithacus Software Ltd.) using an equation modelling a simple two-state denaturation as described previously.<sup>8</sup> Unfolding measured by this method commenced at 0.7 M urea for the wild-type protein (black curve, open circles) and 1.2 M urea for Thr114Phe  $\alpha_1$ -antitrypsin (blue curve, filled circles). The transition midpoints (assuming complete unfolding at 8 M urea) were reached at similar points (3.4 M urea for wild-type  $\alpha_1$ -antitrypsin and 3.5 M for Thr114Phe  $\alpha_1$ -antitrypsin) in both profiles. (d) Electron density for the Thr114Phe mutation site (boxed) prior to restrained refinement (upper panel). The 2.0-Å structure of wild-type  $\alpha_1$ -antitrypsin (1QLP with the mutation site modelled as an alanine) was used as the search model, giving the  $2F_o - F_c$  electron density map shown in dark blue (contoured at 1  $\sigma$ ). Positive density for the phenylalanine residue appears in the  $F_o - F_c$  difference map shown in red (contoured at 3  $\sigma$ ). The same site is shown in the refined structure (lower panel). (e) The surface-accessible cavity flanking  $\beta$ -sheet A, shown in close up for wild-type (left panel, yellow) and in the refined structure of Thr114Phe (middle panel, cyan)  $\alpha_1$ -antitrypsin with side chains and electrostatic surfaces. The volume occupied by the Thr114Phe mutation within the cavity defines a pharmacophore target for screening of fragment compounds to mimic its effects [right panel, cyan mesh, within the wild-type  $\alpha_1$ -antitrypsin cavity (transparent surface representation)]. (f) Results of proof-of-principle fragment screen. Left panel: Ensemble of the 65 highest-scoring ligands successfully docked using the induced fit protocol. Right panel: 4-D plot showing characteristics of these ligands (results shown correspond to 903 poses). All are highly drug-like and favourable for drug development according to the Rule of 5<sup>17</sup> and the Rule of 3.<sup>18</sup> Heat map colouring indicates ligand–target centroid proximity ( $\geq 5.0$  Å, yellow; 2.5–5.0 Å, orange;  $\leq 2.5$  Å, red). Glide docking score values  $< -6.8$  kcal/mol correspond to micromolar binding constants, and values below  $-8.2$  kcal/mol are consistent with predicted submicromolar  $K_d$ . Closer proximity to the pharmacophore target site indicates increased likelihood of a docked moiety mimicking effects of the Thr114Phe mutation. Smaller ligands allow more extensive structural optimisation during drug design.<sup>19</sup> Four ligands that formed an outlier group in terms of particularly scoring using both Glide SP and Prime scoring algorithms are indicated (box). (g) Examples of ligands with favourable characteristics following induced fit docking. The pharmacophore target is represented as a transparent cyan surface, and ‘receptor residues’ (Asn104, Thr114 and His139) are shown as optimised for each ligand. Ligand colours correspond to those of receptor site residues optimised for their binding. Non-carbon atoms are individually coloured (oxygen, red; nitrogen, dark blue; fluorine, light blue; sulfur, orange). Left panel: the five ligands top-ranked by Glide SP docking score. Middle panel: the five ligands top-ranked by proximity of ligand to the pharmacophore target. Right panel: the ‘favourable outlier group’ defined in (f).



highlighted by *in silico* screens that have identified lead compounds that could fill the cavity and, hence, reduce polymerisation.<sup>15</sup> However, unlike the mutations, all compounds that have proven effective *in vitro* have also abolished inhibitory activity.

To understand the structural mechanisms by which the effects of the Thr114Phe and Gly117Phe mutations are conferred, we have crystallised both  $\alpha_1$ -antitrypsin variants and solved their structures to 2.2 and 3.2 Å, resolution, respectively. The structure of Thr114Phe  $\alpha_1$ -antitrypsin allows us to define a novel target for pharmacophores to block polymer formation whilst preserving inhibitory activity. The structure of the Gly117Phe mutant shows that this mutation exerts its effects by changes in both local packing and the position of the F helix. We have gone on to demonstrate that the two mutations ameliorate the defect in folding of Z  $\alpha_1$ -antitrypsin to the native, functional conformation in a COS-7 cell secretion model. These features extend our understanding of the balance between fold stability and metastability in  $\alpha_1$ -antitrypsin. The data are consistent with both proposed  $\beta$ -sheet-A-linked mechanisms of polymerisation. However, in the case of the single-strand linkage model, they can be integrated with previous findings to propose a coordinated mechanism of  $\beta$ -sheet A expansion from the native monomeric protein.

## Results and Discussion

### Effects of the Thr114Phe mutation arise from partial cavity blockade

The 2.2 Å crystal structure of Thr114Phe  $\alpha_1$ -antitrypsin adopts the classic structure of a serpin fold in the native conformation (Fig. 1a and Table 1). The polypeptide chain spanning residues 24–393 is highly ordered with the exception of residues 344–351. These comprise the P15–P8 section of the reactive loop, a section that is disordered or associated with high *B*-factors in other crystal structures of metastable serpins.<sup>9,16,20–22</sup> The key P1 methionine residue that defines the inhibitory specificity of a serpin adopts a canonical conformation.

Comparison of the structure of Thr114Phe  $\alpha_1$ -antitrypsin with the 2.0-Å structure of wild-type  $\alpha_1$ -antitrypsin that is also in the native conformation [Protein Data Bank (PDB) code: 1QLP]<sup>16</sup> shows a high degree of overall similarity between the two proteins (Fig. 1b). The root-mean-square deviation (r.m.s.d.) of the mutant and wild-type structures, based on C $\alpha$  positions, is 0.4 Å. A number of critical regulators of conformational change in  $\alpha_1$ -antitrypsin have previously been characterised. Apart from the cavity flanking  $\beta$ -sheet A, these are the proximal hinge of the reactive loop,<sup>23</sup> the arginine pocket stabilising the loop at the P5 position,<sup>24</sup> interstrand spacing in  $\beta$ -sheet A,<sup>25</sup> strand 1 of  $\beta$ -sheet C,<sup>26</sup> the F helix<sup>9,12,27,28</sup> and the shutter region<sup>29</sup> (Fig. 1a). These are essentially unchanged in the mutant structure compared to that of wild-type  $\alpha_1$ -antitrypsin.

**Table 1.** X-ray data collection and refinement statistics for structures of Thr114Phe  $\alpha_1$ -antitrypsin and Gly117Phe  $\alpha_1$ -antitrypsin

	Thr114Phe $\alpha_1$ -antitrypsin	Gly117Phe $\alpha_1$ -antitrypsin
<i>Data collection</i>		
Cell dimensions		
<i>a</i> (Å)	113.8	58.8
<i>b</i> (Å)	39.6	149.3
<i>c</i> (Å)	90.3	77.3
$\beta$ (°)	105.1	94.11
Molecules in asymmetric unit	1	3
Space group	C2	P2 <sub>1</sub>
Resolution range (Å)	87.0–2.2	77.1–3.2
<i>R</i> <sub>sym</sub> (%)	10.6 (49.3)	13.8 (58.8)
Completeness (%)	94.2 (90.2)	99.7 (99.7)
<i>I</i> / $\sigma$ <i>I</i>	15.3 (5.5)	11.5 (2.5)
Multiplicity	6.8 (6.8)	4.7 (4.8)
<i>Structure refinement</i>		
No. of atoms (protein)	2847	8562
No. of atoms (water)	164	0
<i>R</i> <sub>free</sub> (%)	27.2	29.0
<i>R</i> <sub>cryst</sub> (%)	22.4	24.2
<i>B</i> <sub>ave</sub>		
Main chain	31.8	68.3
Side chain	34.7	74.6
Ramachandran plot statistics (%)		
Favoured regions	89.8	74.6
Allowed regions	9.0	21.7
Generously allowed	1.2	2.8
Disallowed	0.0	0.9
r.m.s.d. from ideality		
Bonds (Å)	0.007	0.010
Angles (°)	1.3	1.6

Values in parentheses refer to the highest-resolution shell.

$R_{\text{sym}} = \sum_i (\sum_j |I_{ij} - \langle I_i \rangle|) / \sum_i \langle I_i \rangle$ , where *j* are the set of observations for each reflection *i*.

$R_{\text{cryst}} = \sum_i ||F_o| - |F_c|| / \sum_i |F_o|$ .

$R_{\text{free}} = R_{\text{cryst}}$  for 5% of reflections omitted from refinement.

*B*<sub>ave</sub> values are average temperature factors for all molecules in the asymmetric unit.

The lack of perturbation elsewhere in the overall structure of Thr114Phe  $\alpha_1$ -antitrypsin implies that the mutation must be mediating its effect on the fold stability of native  $\alpha_1$ -antitrypsin (Fig. 1c) from within the cavity flanking  $\beta$ -sheet A. The cavity is well defined in this high-resolution structure, permitting interpretation of the effect of the mutation on cavity filling at the side-chain level. The precise location of the aromatic side chain of Phe114 was strongly indicated by the initial density seen within the cavity adjacent to residue 114 when this was modelled as an alanine (to minimise model bias) prior to any restrained refinement (Fig. 1d). Phe114 makes no interactions with nearby residues, and therefore, its effects are entirely due to partial filling of the cavity and changes in its packing environment during unfolding and polymerisation. The aromatic group of Phe114 closes the upper part of the cavity (Fig. 1d) with its side chain just 3.7 Å from Leu103 and 4.0 Å from Asn104 at the top of helix D, distances that permit favourable van der Waals contacts. These data support the importance of  $\beta$ -sheet A compaction in stabilising the native conformer against initial unfolding.

### Proof-of-principle *in silico* fragment screening based upon the Thr114Phe mutation

The volume partially filling the cavity in Thr114Phe  $\alpha_1$ -antitrypsin (Fig. 1e) is in a region of high-solvent accessibility in the wild-type protein and also yields the therapeutically desirable combination of effective blockade of polymerisation whilst maintaining inhibitory function. Furthermore, it does not alter the native  $\alpha_1$ -antitrypsin structure beyond the cavity in any way. It is, therefore, attractive to consider it a refined target or an ideal pharmacophore volume, depicted in Fig. 1e as the difference between surface representations of wild-type and Thr114Phe  $\alpha_1$ -antitrypsin around residue 114. The estimated size of this ideal pharmacophore is 59 Å<sup>3</sup>, equivalent to 22% of the total cavity volume in wild-type  $\alpha_1$ -antitrypsin.<sup>16</sup> A similar reduction in the size of lead compounds compared to those identified by screening against the entire cavity would improve the probability of success in rational drug design.<sup>18</sup> Lipinski's 'Rule of 5'<sup>17</sup> observes that success in drug development is highly correlated with the following:  $M_r < 500$  Da, lipophilic index ( $\log P$ )  $< 5$ ,  $\leq 5$  hydrogen-bond donors,  $\leq 10$  hydrogen-bond acceptors. A modified version of this, the 'Rule of 3', has been proposed to guide drug design via exploration of a greater diversity of chemical space using still smaller ( $< 300$  Da) molecular fragments to identify potential 'hits'.<sup>19</sup> The mean change in  $M_r$  in development from lead compound to drug is +69 Da, and most effective drugs used in clinical practice are 250–320 Da.<sup>19</sup> Currently, the lead compounds designed to block the entire cavity and shown to be effective *in vitro* are at the upper end of the desirable drug size (310–650 Da).<sup>15</sup>

To consider the region occupied by the Thr114Phe mutation a meaningful pharmacophore target, it must generate hits *in silico* that are plausible in terms of mimicking the mutation's partial cavity blockade, their predicted binding affinities ( $K_d$ ) and 'drug-likeness'. The pharmacophore target is centred between the middle of strand 2 of  $\beta$ -sheet A (s2A) and the top of helix D (Fig. 1f). It is close to polar residues (Asn104, Thr114 and His139) and an area favourable for hydrophobic contacts that may aid ligand coordination and binding, respectively. On the other hand, unlike the entire cavity, this region is only bounded by residues in one plane, which may be challenging for some docking strategies. These considerations make it hard to assume *a priori* that plausible docking candidates would emerge from *in silico* screening of fragment libraries.

We have therefore performed a proof-of-concept screen of  $> 70,000$  fragments with  $M_r = 150$ –250 Da to assess the potential of the new pharmacophore for identifying small compounds that can be developed in rational drug design. The pharmacophore centroid was set as the midpoint between Thr114 (mutation site, middle of s2A) and Asn104 (top of helix D) for the purposes of our protocol. Screens were performed to select hits on the basis of

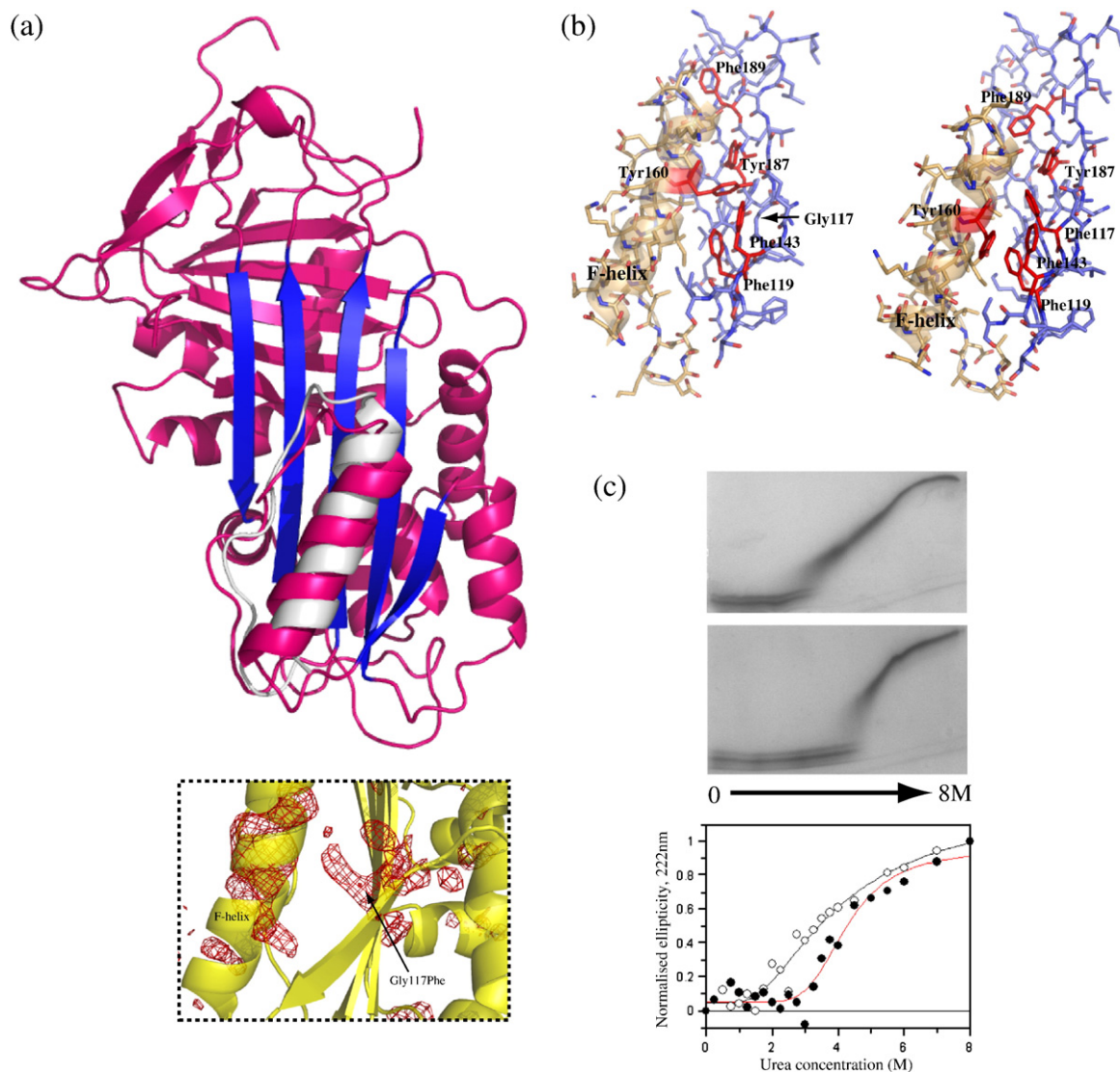
docking scores and ligand–target centroid proximity. By definition, all hits obey the Rule of 5 and can only deviate from the Rule of 3 a single hydrogen-bond acceptor. The best-scoring 0.1% of ligands were taken forward for induced fit docking, resulting in optimised docking of 65 unique ligands. These had the following mean ( $\pm$ SD) characteristics:  $M_r = 207.2$  Da ( $\pm 21.7$ ), estimated  $\log P$  value = 1.76 ( $\pm 0.81$ ) and ligand–pharmacophore centroid proximity = 5.16 Å ( $\pm 0.76$ ). The fragments are shown docked ensemble in relation to residues 104, 114 and 139 (Fig. 1f, left). The best docking scores and proximity to the pharmacophore centroid and sizes of all ligands docked by the induced fitting protocol are also shown (Fig. 1f, right). Docking scores calculated by Glide SP correlate with predicted  $K_d$  values such that a value  $< -6.8$  kcal/mol correlates with a predicted micromolar  $K_d$  and a value  $< -8.2$  kcal/mol correlates with submicromolar affinity. Figure 1f illustrates that this screen identified numerous fragment hits achieving scores within this range. Examples of docked ligands ranked highly by docking scores and fragment–target centroid proximity are shown in Fig. 1g.

Optimisation of this strategy is likely to require screening 1–2 orders of magnitude more fragments and 'growing' (up to  $M_r \sim 250$  Da) of the best hits to optimise specificity and drug-likeness ahead of *in vitro* testing. However, the hits identified by this preliminary screen are good candidates in terms of size, predicted binding affinities and drug-likeness for further development. Moreover, our screen also identified compounds that docked well to sites above the pharmacophore target site or interacted with the front of  $\beta$ -sheet A (Fig. 1f). Similar hits will aid the 'growing' of compounds for future *in vitro* testing.

### The Gly117Phe mutation alters aromatic ring packing in the interface of $\beta$ -sheet A with the F-helix interface and induces a half-turn downward shift in F-helix position

The structure of Gly117Phe  $\alpha_1$ -antitrypsin was solved to 3.2 Å resolution (Table 1 and Fig. 2a), allowing confident interpretation of the secondary structure of the protein. The asymmetric unit contained three molecules with minimal intermolecular contacts and in an arrangement not suggestive of biologically significant multimerisation. No significant changes are evident in  $\beta$ -sheet A, the shutter region or strand 1 of  $\beta$ -sheet C in Gly117Phe  $\alpha_1$ -antitrypsin relative to the wild-type protein. Electron density was not seen for residues 349–356 (copy A), 348–356 (copy B) and 348–358 (copy C) of the hypermobile reactive site loop.

Before incorporation of the Phe mutation into the model or any restrained refinement had been performed, clear density was apparent at the mutation site, indicating that its side chain was oriented away from  $\beta$ -sheet A towards the F helix where significant changes were seen. To investigate



**Fig. 2.** The 3.2-Å structure of Gly117Phe  $\alpha_1$ -antitrypsin. (a) Crystal structure of Gly117Phe  $\alpha_1$ -antitrypsin (pink;  $\beta$ -sheet A in dark blue). The helix F- $\beta$ -sheet A linker region of the superposed native wild-type protein is shown (white). Box: Initial observation of changes around the mutation site following molecular replacement and rigid-body fitting using the wild-type  $\alpha_1$ -antitrypsin search model (1QLP, shown in yellow). At this stage, difference density (shown in red mesh, contoured at 3  $\sigma$ ) for the mutation site and overlying F helix was calculated by excluding residues 117 and 160 (+surrounding 3.5 Å spheres) from a simulated annealing omit map to minimise model bias. Similar difference density was seen in the F-helix region in all three copies of the molecule; the example of a single copy is shown. (b) Side-chain packing of aromatic side chains in the helix F- $\beta$ -sheet A interface in wild-type (left) and Gly117Phe (right)  $\alpha_1$ -antitrypsin. Only  $\beta$ -sheet A (blue, stick representation) and helix F together with its linker region (gold, transparent cartoon and stick representation) are shown. Residues involved in aromatic ring interactions are coloured red, and the mutation site at Gly117 is indicated for wild-type  $\alpha_1$ -antitrypsin. (d) Increased biochemical stability of the native fold in the presence of the Gly117Phe mutation. TUG-PAGE of wild-type (upper panel) and Gly117Phe (middle panel)  $\alpha_1$ -antitrypsin indicates that the mutation confers increased stability of the native fold to increasing urea concentration. This is confirmed by CD spectroscopy (lower panel). For Gly117Phe  $\alpha_1$ -antitrypsin (red curve, filled circles), unfolding begins at 2.0 M urea (compared with 0.7 M for the wild-type protein, black curve, open circles) and the transition midpoint occurs at 4.3 M urea (compared with 3.4 M for wild-type  $\alpha_1$ -antitrypsin). Conditions are as described for Fig. 1c. Taken together, the urea unfolding studies indicate increasing stability of the native fold in the order wild-type < Thr114Phe < Gly117Phe  $\alpha_1$ -antitrypsin.

these changes, we omitted the mutation site and the overlying part of the F helix to generate a difference density map of these areas with model bias minimised. This was based upon the initial rigid-body fit of the search model and did not use noncrystallographic symmetry (NCS) restraints, yet

it confirmed a half-turn shift of the F-helix along its axis for all three molecules as well as side-chain density for the underlying Phe117 residue (Fig. 2a, boxed). In the refined structure, the F-helix is displaced downwards (towards the bottom of s3A) by half a turn relative to its position in the structure

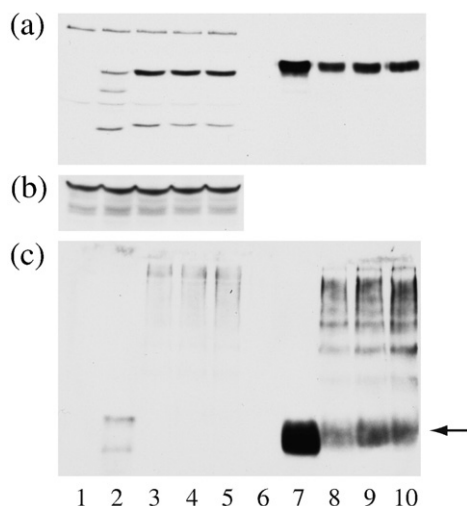


of native wild-type  $\alpha_1$ -antitrypsin (2.7–3.5 Å between equivalent C $^\alpha$  atoms). This compares with an overall C $^\alpha$  r.m.s.d. of 1.2 Å for the rest of the two structures. The F-helix shift is unlikely to be an artefact of crystal packing as different copies of Gly117Phe  $\alpha_1$ -antitrypsin have different packing environments. Moreover, no F-helix changes are required to permit packing of three molecules of the native wild-type protein in orientations identical with those of the three copies of Gly117Phe  $\alpha_1$ -antitrypsin in the asymmetric unit (not shown). Conversely, repacking of helix F– $\beta$ -sheet A interface residues, in particular Tyr160, is necessary to avoid steric clashes with the adjacent Phe117 side chain and sustain favourable hydrophobic interactions. Figure 2b compares packing of the helix F– $\beta$ -sheet A interface in wild-type and Gly117Phe  $\alpha_1$ -antitrypsin. In the refined structure of Gly117Phe  $\alpha_1$ -antitrypsin, Phe117 forms the centre of a cluster of aromatic side chains, interacting with Phe119, Phe143 and Tyr187 (all at distances of 3.5–5.5 Å). The side chain of Tyr160 is in a solvent-exposed position in this mutant in contrast to its buried position within the helix F– $\beta$ -sheet A interface in wild-type  $\alpha_1$ -antitrypsin but remains well placed for aromatic interactions with two residues of this cluster, Phe119 and Phe143. These differences therefore account for the changed conformation of helix F in Gly117Phe  $\alpha_1$ -antitrypsin compared to the wild-type protein. They introduce an extra aromatic ring into the helix F– $\beta$ -sheet A interface in the mutant to increase favourable packing interactions between strands 1–3 of the  $\beta$ -sheet. These changes do not obviously increase the degree of hydrophobic interaction between the F helix (or its linker to s3A, both shown in gold) and  $\beta$ -sheet A (blue). However, it is striking that the biochemistry of this mutant<sup>14</sup> resembles that of  $\alpha_1$ -antitrypsin, carrying what can now be regarded as a complementary mutation, Tyr160Trp.<sup>28</sup> Here, an extra aromatic ring is introduced from the F-helix side into the helix F– $\beta$ -sheet A interface to project into a site adjacent to Gly117. The Tyr160Trp mutation is associated with a near-identical increase in thermal stability to that seen with the Gly117Phe mutation. Both mutations are also alike in achieving this without any reduction in inhibitory activity. The increased biochemical stability conferred by the Gly117Phe mutation on the native fold of  $\alpha_1$ -antitrypsin (Fig. 2c) is therefore related to increased stabilisation of the F helix. The relative contributions of the alterations in packing of the helix F– $\beta$ -sheet A interface and its half-turn displacement towards the lower pole of the molecule remain to be determined. Since these effects involve alterations of the final folding of native  $\alpha_1$ -antitrypsin and are mediated via interactions in a hydrophobic interface, they may be hard to mimic via rational drug design. However, the fact that the position of the F helix is affected in the absence of any other significant changes to the overall fold supports the conclusion that this helix forms relatively late in the folding process and independent from the folding of other structural features.

### Effect of Thr114Phe and Gly117Phe mutations on secretion of Z $\alpha_1$ -antitrypsin from COS-7 cells

To relate our findings to drug design for  $\alpha_1$ -antitrypsin deficiency, we attempted to characterise the stability and folding of the Thr114Phe/Z and Gly117Phe/Z double mutants of  $\alpha_1$ -antitrypsin. These were produced as recombinant proteins using the same overexpression system as the crystallised proteins and purified by the same method. The yield of both double mutants was between 1 and 2 orders of magnitude lower than that of wild-type and single-Phe mutant  $\alpha_1$ -antitrypsin. The products were conformationally heterogeneous with a mixture of native, latent [the major product on non-denaturing and urea polyacrylamide gel electrophoresis (PAGE)] and polymeric conformers. The effects of the Thr114Phe and Gly117Phe mutations on transient expression and secretion of  $\alpha_1$ -antitrypsin containing the Z mutation in a more physiological, eukaryotic cell system were then assessed. Previous work has shown that the Thr114Phe and Gly117Phe mutations partially rescue the profound defect in the overall secretion of  $\alpha_1$ -antitrypsin caused by the Z mutation in a *Xenopus* oocyte system.<sup>14</sup> These effects are modest in proportion to the overall deficit but are of considerable clinical interest since disease severity correlates closely with rates of polymerisation and (inversely) with circulating levels of  $\alpha_1$ -antitrypsin.<sup>30</sup> However, this study did not characterise the conformational state of the secreted material. We therefore transformed and transiently expressed M, Z, Thr114Phe/Z and Gly117Phe/Z  $\alpha_1$ -antitrypsin in COS-7 cells. Analysis of cell lysates and culture media by Western blot analysis of sodium dodecyl sulfate (SDS)-PAGE showed increased secretion of  $\alpha_1$ -antitrypsin in the double mutants compared with Z  $\alpha_1$ -antitrypsin (Fig. 3a and b). These data were consistent with those quantified by the more sensitive pulse-chase system in *Xenopus* oocytes. This increase in total  $\alpha_1$ -antitrypsin secretion was related to a clear increase in secretion of native, monomeric  $\alpha_1$ -antitrypsin as assessed by Western blot analysis of non-denaturing PAGE (Fig. 3c, arrow). The absolute concentrations of  $\alpha_1$ -antitrypsin in the culture media were very low relative to those conventionally used to measure the inhibitory activity of  $\alpha_1$ -antitrypsin. However, we were able to detect inhibitory activity using an ELISA technique in which bovine  $\alpha$ -chymotrypsin was used to capture functional  $\alpha_1$ -antitrypsin from concentrated culture media. This indicated that more functional material was secreted in the presence of the double mutations (1.35  $\mu$ g/ml for Thr114Phe/Z and 0.35  $\mu$ g/ml for Gly117Phe/Z  $\alpha_1$ -antitrypsin) than with the Z mutation alone (<0.16  $\mu$ g/ml of active protein; below linear range of detection by ELISA). These effects were modest compared with the amount of functional M  $\alpha_1$ -antitrypsin detected in the culture media (>5  $\mu$ g/ml; exceeding the maximum assessable in





**Fig. 3.** Characterisation of M, Z, Thr114Phe/Z and Gly117Phe/Z  $\alpha_1$ -antitrypsin secreted from COS-7 cells. Western blot analysis of  $\alpha_1$ -antitrypsin retained within or secreted from COS-7 cells 2 days after transient transfection. (a) Western blot analysis of samples run on SDS-PAGE with (b) luciferase transfection control for cell lysate samples [same gel as (a)]. (c) Western blot analysis of samples run on non-denaturing PAGE. Lanes 1–5, cell lysates; lanes 6–10, culture media. Lane 1, empty vector control—cell lysate; lane 2, M  $\alpha_1$ -antitrypsin—cell lysate; lane 3, Z  $\alpha_1$ -antitrypsin—cell lysate; lane 4, Thr114Phe/Z  $\alpha_1$ -antitrypsin—cell lysate; lane 5, Gly117Phe/Z  $\alpha_1$ -antitrypsin—cell lysate; lane 6, empty vector control—supernatant; lane 7, M  $\alpha_1$ -antitrypsin—supernatant; lane 8, Z  $\alpha_1$ -antitrypsin—supernatant; lane 9, Thr114Phe/Z  $\alpha_1$ -antitrypsin—supernatant; lane 10, Gly117Phe/Z  $\alpha_1$ -antitrypsin—supernatant. Arrow indicates migration of native  $\alpha_1$ -antitrypsin secreted into the culture media on non-denaturing PAGE.

the linear range of the ELISA) but were consistent with the appearances of native protein by non-denaturing PAGE (Fig. 3c).

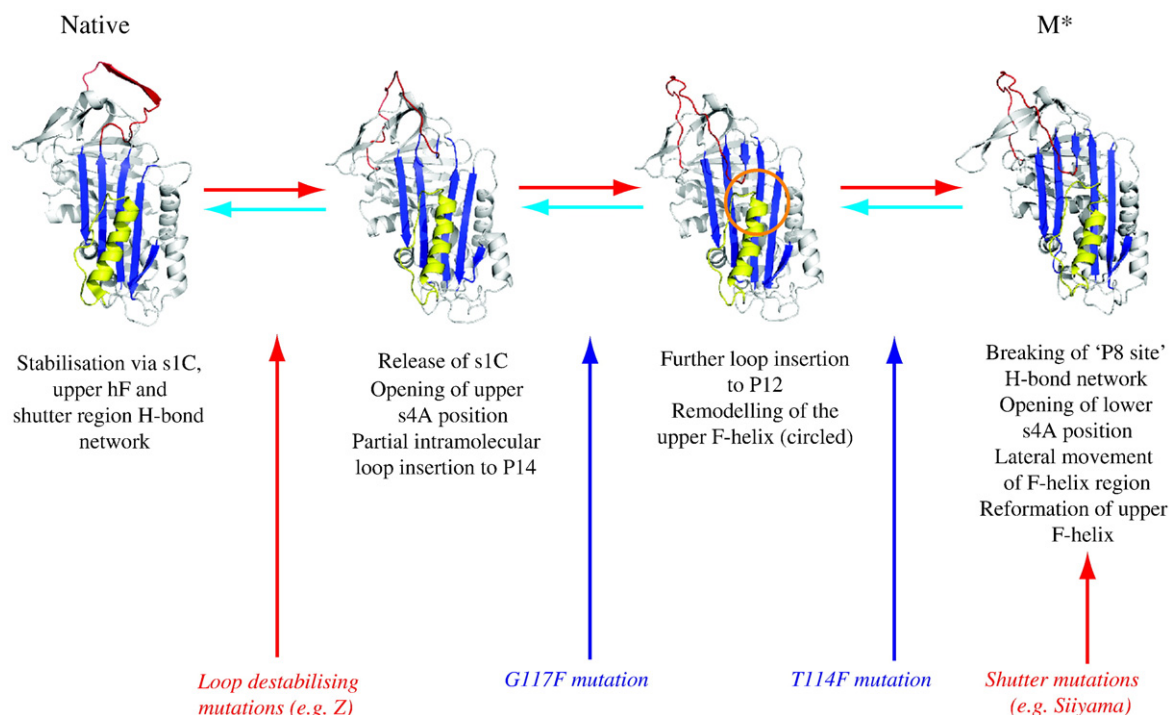
The increased secretion of native  $\alpha_1$ -antitrypsin from cells expressing the double mutants compared to those expressing Z  $\alpha_1$ -antitrypsin was also associated with an increase in the amount of polymeric material seen in the media. We cannot formally exclude the possibility that these polymers were formed intracellularly and then secreted differentially. However, the lack of any observable difference in total and polymeric intracellular  $\alpha_1$ -antitrypsin levels between the different constructs does not support this hypothesis. We therefore favour the alternative explanation that the extracellular polymers are formed from  $\alpha_1$ -antitrypsin that is secreted in the native conformation. Taken together, these data support the conclusion that the Thr114Phe and Gly117Phe mutations have modest but detectable effects on increasing secretion of functional native  $\alpha_1$ -antitrypsin on the background of the severe Z mutation in eukaryotic cells. However, these double mutants remain prone to conformational change and polymerisation.

### Relating the crystal structures of Thr114Phe and Gly117Phe $\alpha_1$ -antitrypsin to the single- and double-strand linkage models of polymerisation

The Thr114Phe and Gly117Phe mutations act by different mechanisms to increase the global stability of native  $\alpha_1$ -antitrypsin. Expansion of  $\beta$ -sheet A is a common feature in both proposed mechanisms of polymerisation.<sup>9,13</sup> In order for  $\alpha_1$ -antitrypsin to adopt either M\* conformation, residue 114 must move laterally by 4.6 Å in the plane of  $\beta$ -sheet A. This is limited by local steric clashes in Thr114Phe  $\alpha_1$ -antitrypsin. In the case of the Gly117Phe mutation, the structural mechanism whereby the changes in the vicinity of the F helix may impede formation of a  $\beta$ -hairpin donor/acceptor M\* species is less intuitive. There is no apparent increase in stabilising interactions between the helix F–s3A linker and s5A, and the changes in packing in the interface between the F helix and  $\beta$ -sheet A would not of themselves be expected to impede opening of the  $\beta$ -sheet. Nevertheless, our data are not inconsistent with this model since they may simply reflect overall cooperativity of the final steps on the folding pathway (and initial steps on the unfolding pathway) for  $\alpha_1$ -antitrypsin. However, our findings can be mechanistically integrated with other biochemical data for the native-to-M\* transition posited for the single-strand linkage model in which opening of the s4A site is associated with partial intramolecular loop insertion. Various data support the following features of this model: upper s4A opening precedes lower s4A opening<sup>9,25,31,32</sup> and is associated with partial insertion of the reactive site loop.<sup>9,33,34</sup> Subsequent lower s4A opening around the site of P8 residue insertion is associated with breaking of a network of interactions between s4A, s5A and shutter region residues.<sup>25</sup> This can be induced by the insertion of a cleaved reactive loop or a peptide analogue into  $\beta$ -sheet A from the P14–P9 sites, producing species that are highly polymerogenic.<sup>31,32,35,36</sup> In  $\alpha_1$ -antitrypsin, modelling the insertion of the reactive site loop to the P12 position necessitates both the release of s1C and remodelling of the upper helix F—secondary structural changes known to occur during the formation of polymers from native  $\alpha_1$ -antitrypsin.<sup>12,26</sup> However, stabilisation of the upper turns of the F helix and/or its downward displacement as seen in the structure of Gly117Phe  $\alpha_1$ -antitrypsin will reduce the propensity for its remodelling. It will, therefore, reduce M\* formation if such remodelling is integral to the formation of the intermediate. This process is shown as a schema in Fig. 4. The model of the single-strand donor/acceptor M\* is shown following energy minimisation and simulated annealing molecular dynamics (coordinates are supplied as [Supplementary Data](#)).

### Conclusion

The data presented here demonstrate in crystallographic detail two mechanisms that greatly stabilise the native fold in  $\alpha_1$ -antitrypsin without abolishing its inhibitory function. These mechanisms are likely to



**Fig. 4.** Schema for formation of a partially loop-inserted M\* species from native  $\alpha_1$ -antitrypsin. Findings from the Thr114Phe and Gly117Phe  $\alpha_1$ -antitrypsin crystal structures are incorporated together with previous observations of requirements for loss of the s1C<sup>26</sup> strand, remodeling of the F helix<sup>9,12,28</sup> and destabilising of interactions involving shutter region residues.<sup>25</sup> Sequential insertion of the reactive loop into the upper s4A position is depicted by two chimeras. These are derived from structures of native (1QLP) and latent (1IZ2)  $\alpha_1$ -antitrypsin, murine  $\alpha_1$ -antichymotrypsin (1YXA—demonstrating changes associated with opening of the P14 acceptor site) and thyroxine binding globulin (2CEO—demonstrating changes on expansion at the P12 insertion site). In all cases, the reactive loop modelled is that of  $\alpha_1$ -antitrypsin. The final image shows the  $\alpha_1$ -antitrypsin M\* model, generated as described in Materials and Methods, after energy minimisation with simulated annealing molecular dynamics. The effects of mutations that will facilitate this transition are shown in red whilst those that block it are shown in blue. Thus, in this scheme, Thr114Phe retards M\* formation by cushioning  $\beta$ -sheet A against expansion whilst Gly117Phe stabilises the upper turns of helix F and abolishes the steric requirement for their remodelling in response to partial loop insertion. Conversely, mutations may accelerate polymerisation by favouring partial loop insertion (e.g., Z) or by opening the lower s4A position directly (e.g., Siiyama).

underlie the partial rescue of secretion of native  $\alpha_1$ -antitrypsin when the Thr114Phe or Gly117Phe mutations occur in the context of the Z mutation in eukaryotic cells although the double mutants remain highly polymerogenic. The high-resolution structural characterisation of the Thr114Phe mutation presents a more precise pharmacophore target for screening of small fragments capable of mimicking its effects. These data are generally consistent with either the single-strand or  $\beta$ -hairpin linkage models of polymerisation. However, they suggest a specific mechanism of propagated conformational change in the case of the single-strand model and, therefore, open up new approaches for further studies to test its applicability in this important pathological pathway.

## Materials and Methods

### Site-directed mutagenesis, protein expression, purification and characterisation

QuikChange mutagenesis (Stratagene, California, USA) was used to introduce the Thr114Phe, Gly117Phe and/or Z (Glu342Lys) mutations into His-tagged  $\alpha_1$ -antitrypsin in

the pQE31 vector<sup>37</sup> to create single- or double-mutant constructs. The full-length sequences of the mutated  $\alpha_1$ -antitrypsin genes were confirmed, and plasmids were transformed into XL-1 Blue *Escherichia coli* cells. The wild-type and mutant proteins were expressed and purified as described previously.<sup>14</sup> The final products were stored in 10 mM Tris/HCl, pH 7.4, 10 mM NaCl and 1 mM ethylenediaminetetraacetic acid. Product purities were confirmed by non-denaturing PAGE, SDS-PAGE and transverse urea gradient (TUG)-PAGE. Further characterisation of urea stability by circular dichroism (CD) was performed as described previously.

The Z, Thr114Phe/Z and Gly117Phe/Z  $\alpha_1$ -antitrypsin constructs for transient transfection of COS-7 cells were prepared by a combination of cloning, subcloning and site-directed mutagenesis in the pcDNA3.1 vector. All constructs were confirmed by sequencing of the whole gene. COS-7 cells were transfected and cultured as described previously.<sup>38</sup> Samples of the cells in culture media were collected 2 days after transfection. Cells and culture media were separated by centrifugation at 500g for 10 min. Cell lysates and culture media were prepared and assessed by SDS-PAGE and non-denaturing PAGE followed by Western blot analysis for  $\alpha_1$ -antitrypsin as described previously.<sup>38,39</sup> Inhibitory activity was assessed with an  $\alpha$ -chymotrypsin-based ELISA. Each well of a 96-well plate was coated overnight with 100  $\mu$ l of 5 mg/ml

bovine  $\alpha$ -chymotrypsin. The plate was then washed before incubating for 15 min at 37 °C with 100  $\mu$ l of culture medium concentrated 7.5 times (this did not affect relative proportions of polymeric to total  $\alpha_1$ -antitrypsin). After further washing to remove unbound material, the amount of functional material captured was assessed by incubation with 100  $\mu$ l of purified polyclonal antibody against  $\alpha_1$ -antitrypsin used at 1  $\mu$ g/ml for 2 h at 4 °C in the presence of 0.25% (w/v) bovine serum albumin as a blocking agent, followed by washing and incubation with 100  $\mu$ l of horseradish-peroxidase-conjugated anti-rabbit secondary antibody (Sigma-Aldrich, Dorset, UK). The plate was washed and the enzymatic activity was developed with 3,3',5,5'-tetramethylbenzidine substrate (Sigma-Aldrich). The colour was allowed to develop for approximately 10 min and was halted by the addition of 100  $\mu$ l of 1 M sulfuric acid to each well. The absorbance was read at 450 nm.

### Crystallisation, data collection and structure solution

Both Thr114Phe and Gly117Phe  $\alpha_1$ -antitrypsin were crystallised by the hanging drop vapour diffusion method in 4- $\mu$ l drops containing equivalent volumes of 10 mg/ml protein and mother liquor at 20 °C. Thr114Phe  $\alpha_1$ -antitrypsin was crystallised using 0.1 M  $\text{Li}_2\text{SO}_4$ , 0.1 M Tris/HCl pH 8.5, and 31% polyethylene glycol (PEG) 4000 (w/v). Gly117Phe  $\alpha_1$ -antitrypsin was crystallised using 0.1 M  $\text{Li}_2\text{SO}_4$ , 0.1 M Tris/HCl, pH 8.5, and 25% PEG 4000 (w/v). Crystals were cryoprotected by the addition of 16% PEG 400 (w/v) (for Thr114Phe  $\alpha_1$ -antitrypsin) or 16% ethylene glycol (w/v) (for Gly117Phe  $\alpha_1$ -antitrypsin) before being flash cooled in liquid nitrogen. Diffraction data for Thr114Phe  $\alpha_1$ -antitrypsin were collected at station I03 at the Diamond synchrotron (Oxfordshire, UK), and those for Gly117Phe  $\alpha_1$ -antitrypsin were collected at station ID14.4 at the European Synchrotron Radiation Facility (Grenoble, France). Data were processed using MOSFLM<sup>40</sup> and scaled using Scala.<sup>41</sup> Molecular replacement was carried out using MOLREP<sup>42</sup> with protein coordinates from the 2.0-Å structure of wild-type  $\alpha_1$ -antitrypsin (1QLP)<sup>16</sup> as the search model. Refinements for both structures were carried out using Refmac<sup>43</sup> and CNS.<sup>44</sup> Whilst the asymmetric unit of Thr114Phe  $\alpha_1$ -antitrypsin contained only a single molecule, that of the Gly117Phe mutant contained three molecules. Two molecules were related by a 2-fold rotational symmetry axis, the third was related to the others by translation without an  $n$ -fold rotational axis. NCS was therefore applied during the solution of the structure of Gly117Phe  $\alpha_1$ -antitrypsin. This was initially applied with all three molecules constrained to be identical. Once  $R_{\text{free}}$  values had converged between successive refinement rounds, the NCS was subsequently applied with weighted restraints. In three regions where the electron density indicated increased intermolecular variability [residues 90–112 (helix D and linker to s2A), 150–180 (helix F and linker to s3A) and 343–363 (reactive site loop)], NCS restraints were not applied. Manual rebuilding of models was performed in Coot<sup>45</sup> and figures were produced using PyMOL<sup>46</sup> and Chimera.<sup>47</sup>

### Characterising the pharmacophore target implied by the Thr114Phe mutation

The cavity flanking  $\beta$ -sheet A in  $\alpha_1$ -antitrypsin was defined and quantified using SURFNET.<sup>48</sup> As described previously,<sup>16</sup> grid separations of 0.8 Å and probe spheres of 2.0–4.0 Å radii were used. The extra volume occupied within the cavity by the Thr114Phe mutation was

calculated by subtracting the volume of the cavity in the mutant from that in the wild-type protein. Similarly, an ideal pharmacophore capable of mimicking the partial cavity blockade caused by the mutation was illustrated by subtraction of the surface map of residue 114 in wild-type  $\alpha_1$ -antitrypsin from that of Thr114Phe  $\alpha_1$ -antitrypsin using SURFNET.<sup>48</sup>

### In silico screens of compound libraries

The fragment-like subset of the ZINC database<sup>49</sup> ( $-2 \leq \text{xLogP} \leq 3$ ,  $M_r = 150\text{--}250$  Da,  $\leq 3$  rotational bonds,  $\leq 2$  hydrogen-bond donors and  $\leq 4$  hydrogen-bond acceptors) was selected for initial prescreening to identify drug- and lead-like compounds capable of significant modification to optimise specificity without unfavourably increasing size. The subset labelled 'usual', containing all fragment-like molecules protonated in the pH range 5.75 to 8.25, was downloaded. Programs from the Schrödinger Suite 2007 (Schrödinger LLP, New York, NY, USA) were used for *in silico* docking. Fragments were preprocessed using LigPrep version 2.1. The original ionisation state was retained and 'desalt' and 'tautomer generation' flag settings were used. Chiralities were determined from the original ZINC database structures, and one low-energy ring conformation was allowed. This procedure resulted in a total of 70,925 small-molecule structures for docking. The protein was prepared using Maestro version 8.0 (Protein Preparation Wizard). Hydrogens were added, and all waters were deleted from the structure of wild-type  $\alpha_1$ -antitrypsin. Hydrogen bonds were optimised, and the structure was minimised within 0.30 Å r.m.s.d. from the initial structure. The potential energy grid was pre-calculated using a box centred on the centroid of residues Asn104 and Thr114 to prepare the binding site for docking. The box was 39 Å in each dimension, but the ligand centre was constrained to lie within 14 Å of the box centre. No additional constraints were used.

Initial screening of the 70,925 ligands was performed using Glide version 4.5, to generate a range of putative ligand–receptor interactions (poses). The best pose for each ligand was calculated using Emodel and retained. Ranking between ligands was then performed using the Glide SP score. All Glide parameters were left to their default values. The top 70 (<0.1%) ligands resulting from this Glide SP run were then submitted to the Induced Fit Docking protocol, a three-step docking process. Preliminary docking by Glide SP used a scaling factor for van der Waals radii of 0.5 for both the ligand and the receptor. For each ligand, the best 20 poses were retained. Prime version 1.6 was used to optimise receptor side chains within 5.0 Å of each posed ligand. Finally, ligands were redocked into the optimised structures. The resulting poses were ranked using a composite of the Glide SP (binding site energy) and the Prime (molecular mechanics + solvation energy function for the protein–ligand complex) scores as previously described.<sup>50</sup>

The molecular weight and other descriptors for the set of docked ligands were calculated using the MOE suite (MOE 2006.08, Chemical Computing Group, Montreal, QC, Canada). Prior to descriptor calculation, the partial charges on all atoms were recalculated using the default MOE force field (MMFF94x), and adjustment of hydrogens and lone pairs was allowed.

### Modelling $\alpha_1$ -antitrypsin in the M\* conformation

To model wild-type and mutant  $\alpha_1$ -antitrypsin in the partially loop-inserted M\* conformation, we used the 2.2-



Å crystal structure of latent  $\alpha_1$ -antitrypsin (1QLP)<sup>51</sup> as the initial template. In the latent conformation, the reactive loop is fully inserted into  $\beta$ -sheet A as its fourth strand (s4A). This structure was chosen as it is the highest-resolution structure of  $\alpha_1$ -antitrypsin with an expanded  $\beta$ -sheet A. Mutations were restored back to the wild-type sequence. The residues of the reactive site loop were threaded into the conformation adopted in the structure of  $\delta$  Leu55Pro  $\alpha_1$ -antichymotrypsin (1QMN)<sup>9</sup> using SWISS-MODEL.<sup>52</sup> As the reactive loop of  $\alpha_1$ -antitrypsin is four residues shorter than that of  $\alpha_1$ -antichymotrypsin, this required stripping out of s1C to connect it directly with s4B. The model was regularised in Coot before undergoing energy minimisation, simulated annealing and final energy minimisation using SANDER, with the protein parameterised using the AMBER99SB force field.<sup>53</sup> The upper turn of helix F required greater freedom during simulated annealing than the rest of the molecule to allow successful energy minimisation. This was achieved by conservative relaxation of restraints on this region relative to the rest of the molecule.

### Accession numbers

Coordinates and structure factors are deposited in the PDB with accession number 3DRM for the structure of Thr114Phe  $\alpha_1$ -antitrypsin and 3DRU for the structure of Gly117Phe  $\alpha_1$ -antitrypsin.

### Acknowledgements

This work was supported by a Wellcome Trust Intermediate Clinical Fellowship to B.G., an Alpha1-antitrypsin Laurell Training Award award to M.M., a Wellcome Trust summer studentship to S.C.L.B., a Wellcome Trust Clinical Training Fellowship to R.L.P., the Medical Research Council (UK), the British Pharmaceutical Society and Papworth National Health Service Trust. C.S. is a Wellcome Trust Research Training Fellow. We thank Dr Nora Cronin (School of Crystallography, Birkbeck College, University of London) for her assistance in collecting crystallographic data and Ms Heike Kröger (Cambridge Institute for Medical Research) for her assistance with the cell biological experiments.

### Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2009.01.069](https://doi.org/10.1016/j.jmb.2009.01.069)

### References

- Baugh, R. J. & Travis, J. (1976). Human leukocyte granule elastase: rapid isolation and characterization. *Biochemistry*, **15**, 836–841.
- Silverman, G. A., Bird, P. I., Carrell, R. W., Church, F. C., Coughlin, P. B., Gettins, P. *et al.* (2001). The serpins are an expanding superfamily of structurally similar but functionally diverse proteins. Evolution, novel functions, mechanism of inhibition and a revised nomenclature. *J. Biol. Chem.* **276**, 33293–33296.
- Huber, R. & Carrell, R. W. (1989). Implications of the three-dimensional structure of  $\alpha_1$ -antitrypsin for structure and function of serpins. *Biochemistry*, **28**, 8951–8966.
- Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. 1. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
- Huntington, J. A., Read, R. J. & Carrell, R. W. (2000). Structure of a serpin–protease complex shows inhibition by deformation. *Nature*, **407**, 923–926.
- Mast, A. E., Enghild, J. J., Pizzo, S. V. & Salvesen, G. (1991). Analysis of the plasma elimination kinetics and conformational stabilities of native, proteinase-complexed, and reactive site cleaved serpins: comparison of  $\alpha_1$ -proteinase inhibitor,  $\alpha_1$ -antichymotrypsin, antithrombin III,  $\alpha_2$ -antiplasmin, angiotensinogen, and ovalbumin. *Biochemistry*, **30**, 1723–1730.
- Lomas, D. A., Evans, D. L., Finch, J. T. & Carrell, R. W. (1992). The mechanism of Z  $\alpha_1$ -antitrypsin accumulation in the liver. *Nature*, **357**, 605–607.
- Dafforn, T. R., Mahadeva, R., Elliott, P. R., Sivasothy, P. & Lomas, D. A. (1999). A kinetic mechanism for the polymerisation of  $\alpha_1$ -antitrypsin. *J. Biol. Chem.* **274**, 9548–9555.
- Gooptu, B., Hazes, B., Chang, W.-S. W., Dafforn, T. R., Carrell, R. W., Read, R. & Lomas, D. A. (2000). Inactive conformation of the serpin  $\alpha_1$ -antichymotrypsin indicates two stage insertion of the reactive loop; implications for inhibitory function and conformational disease. *Proc. Natl Acad. Sci. USA*, **97**, 67–72.
- Sivasothy, P., Dafforn, T. R., Gettins, P. G. W. & Lomas, D. A. (2000). Pathogenic  $\alpha_1$ -antitrypsin polymers are formed by reactive loop- $\beta$ -sheet A linkage. *J. Biol. Chem.* **275**, 33663–33668.
- James, E. L. & Bottomley, S. P. (1998). The mechanism of  $\alpha_1$ -antitrypsin polymerization probed by fluorescence spectroscopy. *Arch. Biochem. Biophys.* **356**, 296–300.
- Cabrita, L. D., Dai, W. & Bottomley, S. P. (2004). Different conformational changes within the F-helix occur during serpin folding, polymerization and proteinase inhibition. *Biochemistry*, **43**, 9834–9839.
- Yamasaki, M., Li, W., Johnson, D. J. & Huntington, J. A. (2008). Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature*, **455**, 1255–1258.
- Parfrey, H., Mahadeva, R., Ravenhill, N., Zhou, A., Dafforn, T. R., Foreman, R. C. & Lomas, D. A. (2003). Targeting a surface cavity of  $\alpha_1$ -antitrypsin to prevent conformational disease. *J. Biol. Chem.* **278**, 33060–33066.
- Mallya, M., Phillips, R. L., Saldanha, S. A., Gooptu, B., Leigh Brown, S. C., Termine, D. J. *et al.* (2007). Small molecules block the polymerization of Z  $\alpha_1$ -antitrypsin and increase the clearance of intracellular aggregates. *J. Med. Chem.* **50**, 5357–5363.
- Elliott, P. R., Pei, X. Y., Dafforn, T. R. & Lomas, D. A. (2000). Topography of a 2.0 Å structure of  $\alpha_1$ -antitrypsin reveals targets for rational drug design to prevent conformational disease. *Protein Sci.* **9**, 1274–1281.
- Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **23**, 3–25.
- Carr, R. A. E., Congreve, M., Murray, C. W. & Rees, D. C. (2005). Fragment-based lead discovery: leads by design. *Drug Discovery Today*, **10**, 987–992.

19. Leach, A. R., Hann, M. M., Burrows, J. N. & Griffen, E. J. (2006). Fragment screening: an introduction. *Mol. Biosyst.* **2**, 429–446.
20. Baglin, T. P., Carrell, R. W., Church, F. C., Esmon, C. T. & Huntington, J. A. (2002). Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism. *Proc. Natl Acad. Sci. USA*, **99**, 11079–11084.
21. McGowan, S., Buckle, A. M., Irving, J. A., Ong, P. C., Bashtannyk-Puhalovich, T. A., Kan, W. T. *et al.* (2006). X-ray crystal structure of MENT: evidence for functional loop-sheet polymers in chromatin condensation. *EMBO J.* **25**, 3144–3155.
22. Jin, L., Abrahams, J.-P., Skinner, R., Petitou, M., Pike, R. & Carrell, R. W. (1997). The anticoagulant activation of antithrombin by heparin. *Proc. Natl Acad. Sci. USA*, **94**, 14683–14688.
23. Aulak, K. S., Eldering, E., Hack, C. E., Lubbers, Y. P. T., Harrison, R. A., Mast, A. *et al.* (1993). A hinge region mutation in C1-inhibitor (Ala<sup>436</sup>→Thr) results in nonsubstrate-like behavior and in polymerization of the molecule. *J. Biol. Chem.* **268**, 18088–18094.
24. Elliott, P. R., Abrahams, J.-P. & Lomas, D. A. (1998). Wildtype  $\alpha_1$ -antitrypsin is in the canonical inhibitory conformation. *J. Mol. Biol.* **275**, 419–425.
25. Zhou, A., Stein, P. E., Huntington, J. A. & Carrell, R. W. (2003). Serpin polymerisation is prevented by a hydrogen bond network that is centered on His-334 and stabilized by glycerol. *J. Biol. Chem.* **278**, 15116–15122.
26. Chang, W.-S. W., Whisstock, J., Hopkins, P. C. R., Lesk, A. M., Carrell, R. W. & Wardell, M. R. (1997). Importance of the release of strand 1C to the polymerisation mechanism of serpins. *Protein Sci.* **6**, 89–98.
27. Tsutsui, Y., Lu, L., Gershenson, A. & Wintrobe, P. L. (2006). The conformational dynamics of a metastable serpin studied by hydrogen exchange and mass spectrometry. *Biochemistry*, **45**, 6561–6569.
28. Cabrita, L. D., Whisstock, J. C. & Bottomley, S. P. (2002). Probing the role of the F-helix in serpin stability through a single tryptophan substitution. *Biochemistry*, **41**, 4575–4581.
29. Stein, P. E. & Carrell, R. W. (1995). What do dysfunctional serpins tell us about molecular mobility and disease? *Nat. Struct. Biol.* **2**, 96–113.
30. Mahadeva, R., Chang, W.-S. W., Dafforn, T., Oakley, D. J., Foreman, R. C., Calvin, J. *et al.* (1999). Heteropolymerisation of S, I and Z  $\alpha_1$ -antitrypsin and liver cirrhosis. *J. Clin. Invest.* **103**, 999–1006.
31. Fitton, H. L., Pike, R. N., Carrell, R. W. & Chang, W.-S. W. (1997). Mechanisms of antithrombin polymerisation and heparin activation probed by insertion of synthetic reactive loop peptides. *Biol. Chem.* **378**, 1059–1063.
32. Huntington, J. A., Pannu, N. S., Hazes, B., Read, R., Lomas, D. A. & Carrell, R. W. (1999). A 2.6 Å structure of a serpin polymer and implications for conformational disease. *J. Mol. Biol.* **293**, 449–455.
33. Mahadeva, R., Dafforn, T. R., Carrell, R. W. & Lomas, D. A. (2002). Six-mer peptide selectively anneals to a pathogenic serpin conformation and blocks polymerisation: implications for the prevention of Z  $\alpha_1$ -antitrypsin related cirrhosis. *J. Biol. Chem.* **277**, 6771–6774.
34. Eldering, E., Verpy, E., Roem, D., Meo, T. & Tosi, M. (1995). COOH-terminal substitutions in the serpin C1 inhibitor that cause loop overinsertion and subsequent multimerization. *J. Biol. Chem.* **270**, 2579–2587.
35. Zhou, A., Stein, P. E., Huntington, J. A., Sivasothy, P., Lomas, D. A. & Carrell, R. W. (2004). How small peptides block and reverse serpin polymerization. *J. Mol. Biol.* **342**, 931–941.
36. Dunstone, M. A., Dai, W., Whisstock, J. C., Rossjohn, J., Pike, R. N., Feil, S. C. *et al.* (2000). Cleaved antitrypsin polymers at atomic resolution. *Protein Sci.* **9**, 417–420.
37. Zhou, A., Carrell, R. W. & Huntington, J. A. (2001). The serpin inhibitory mechanism is critically dependent on the length of the reactive center loop. *J. Biol. Chem.* **276**, 27541–27547.
38. Miranda, E., Römisch, K. & Lomas, D. A. (2004). Mutants of neuroserpin that cause dementia accumulate as polymers within the endoplasmic reticulum. *J. Biol. Chem.* **279**, 28283–28291.
39. Miranda, E., MacLeod, I., Davies, M. J., Perez, J., Römisch, K., Crowther, D. C. & Lomas, D. A. (2008). The intracellular accumulation of polymeric neuroserpin explains the severity of the dementia FENIB. *Hum. Mol. Genet.* **17**, 1527–1539.
40. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image data. *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, vol. 26. SERC, Daresbury Laboratory, Warrington, UK.
41. Kabsch, W. (1988). Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J. Appl. Crystallogr.* **21**, 916–924.
42. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025.
43. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
44. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
45. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
46. DeLano, W. L. (2002). The PyMOL Molecular Graphics System. <http://www.pymol.org>.
47. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
48. Laskowski, R. A. (1995). SURFNET: a program for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graphics*, **13**, 323–330.
49. Irwin, J. J. & Schoichet, B. K. (2005). ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182.
50. Sherman, W., Day, T., Jacobson, M. P., Freisner, R. A. & Farid, R. (2006). Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **49**, 534–553.
51. Im, H., Woo, M.-S., Hwang, K. Y. & Yu, M.-H. (2002). Interactions causing the kinetic trap in serpin protein folding. *J. Biol. Chem.* **277**, 46347–46354.
52. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
53. Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E. *et al.* (2006). AMBER 9 University of California, San Francisco, CA.